THE JOURNAL OF ASIA TEFL
Vol. 19, No. 1, Spring, 2022, 17-34
http://dx.doi.org/10.18823/asiatefl.2022.19.1.2.17

The Journal of Asia TEFL
http://journal.asiatefl.org/
e-ISSN 2466-1511 © 2004 AsiaTEFL.org. All rights reserved.

L2 English Knowledge: Implicit, Explicit, or Automatized? The Role of Structure Difficulty

Sepideh Mehraein
University of Tehran, Iran

Hamideh Marefat
University of Tehran, Iran

Hossein Karami
University of Tehran, Iran

The purpose of the current study was to examine the effect structure difficulty may have on the type of knowledge representation L2 learners draw on in different measures of linguistic knowledge. To this aim, word monitoring task (WMT), elicited imitation task (EI), timed grammaticality judgment test (TGJT), untimed grammaticality judgment test (UGJT), and metalinguistic knowledge test (MKT), each consisting of an equal number of easy and difficult structures, were employed. The results of confirmatory factor analysis (CFA) demonstrated that the easy items of GJTs, irrespective of time condition, tap into the construct of automatized knowledge and the difficult ones of EI, TGJT, UGJT, and MKT into explicit knowledge. Additionally, in all the tests the participants' performance on the easy items significantly differed from that on the difficult ones. Therefore, it can be concluded that structure difficulty is a contributing factor in discriminating different knowledge types and has a pivotal role in learners' use of linguistic knowledge.

Keywords: implicit knowledge, explicit knowledge, automatized knowledge, structure difficulty, confirmatory factor analysis

Introduction

Implicit and explicit second language (L2) knowledge are two main concepts in the field of second language acquisition (SLA). Explicit knowledge is the knowledge we are consciously aware of, while implicit knowledge is the knowledge that we are not aware of (DeKeyser, 2009; Hulstijn, 2005). As the ultimate goal of language pedagogy is improving spontaneous language use, that is implicit knowledge, the relationship between these two types of knowledge is of particular importance for SLA research (R. Ellis, 2008). However, the lack of validated measures of these two types of knowledge seems to prevent empirical examination of their relationship. Acknowledging the lack of the measures, R. Ellis (2005, 2009) pioneered attempts to develop measures such as elicited imitation (EI), oral narrative test (ONT), timed grammaticality judgment test (TGJT), untimed grammaticality judgment test (UGJT), and metalinguistic knowledge test (MKT). Different factor-analytic validity studies suggested that EI, ONT, and TGJT serve as measures of implicit knowledge, and UGJT and MKT as measures of explicit knowledge (Bowles, 2011; R. Ellis, 2005; Gutiérrez, 2013; Han & Ellis, 1998; Zhang, 2015).

Although previous studies provided evidence for the construct validity of the above-mentioned measures, the debate on their validity is still present. In measuring implicit and explicit knowledge, the extent to which these two types of knowledge are involved depends on a number of factors including time pressure and task stimulus. A number of previous studies have considered these two important factors (e.g., Bowles, 2011; R. Ellis, 2005; Gutiérrez, 2013; Han & Ellis, 1998; Zhang, 2015; Vafaee et al., 2017), however, the role of structure difficulty (i.e., easy vs. difficult) has received less attention (Bialystok, 1979; Shiu et al., 2018). Moreover, these few studies investigated the effect of structure difficulty through using only one measure, i.e., GJT, not through a test battery of different knowledge types, from more exact measures hypothesized to tap more implicit knowledge such as psycholinguistics ones to measures hypothesized to tap more explicit knowledge such as MKT.

To fill this gap, the present study is after finding out (a) whether WMT, EI, TGJT, UGJT, MKT including easy and difficult items tap into different constructs of implicit, explicit, and/or automatized knowledge, and (b) whether performance on easy/difficult items across the target structures in WMT, EI, TGJT, UGJT, MKT differs significantly. If structure difficulty has a role in tapping different knowledge types, then it is predicted that each will load on a different construct, and there will be a significant difference between learners' performance on the easy and difficult items.

Literature Review

Implicit, Explicit and Automatized Knowledge of Language

Linguistic knowledge consists of implicit and explicit knowledge. Explicit knowledge is defined as knowledge that is represented declaratively, can be verbalized and brought into awareness, whereas implicit knowledge is defined as knowledge that cannot be brought into awareness or articulated (Anderson, 2005; Hulstijn, 2005). In the context of SLA, it is argued that during the fluent use of language, the learners draw on implicit processes and focus their attention on meaning rather than form. On the other hand, explicit processes can draw on knowledge assimilated from external sources (e.g., L2 textbooks or teacher input), or on knowledge internally derived by the learner through reflection on linguistic exemplars (Barrot, 2020; N. Ellis, 2005).

Moreover, differences exist between a third type of knowledge (i.e., automatized knowledge) and implicit knowledge (Suzuki & DeKeyser, 2015; Vafaee et al., 2017). It was formerly assumed that implicit knowledge is drawn on under time pressure, but when the awareness criterion is applied, implicit knowledge should be distinguished from automatized knowledge (Suzuki & DeKeyser, 2015). Jacoby (1991) believes that automaticity is often characterized as lack of awareness. However, because automatization is a long process, automatized knowledge has been defined as a body of conscious knowledge involving different levels of automatization. Automatized knowledge is thus conscious knowledge that is partly automatized, which is distinguished from implicit knowledge. In automatized knowledge, access to knowledge involves the use of language with awareness even though the implementation is rapid. Furthermore, automatized knowledge is not the same as implicit because it results from explicit declarative knowledge. Therefore, automatized knowledge can stand as a distinct construct of knowledge. To wrap it up, automatized and implicit knowledge can be measured as separate constructs.

Measures of Implicit and Explicit Knowledge in SLA

Previous studies have shown that implicit and explicit knowledge are separate constructs and can be measured with different measures (Bowles, 2011; R. Ellis, 2005; Gutiérrez, 2013; Zhang, 2015). Han and Ellis (1998) examined ways of measuring implicit and explicit L2 knowledge. They used a timed oral production test (OPT), a TGJT, a delayed GJT, and a metalingual knowledge interview in order to explore learners' knowledge of verb complementation in English. Forty-eight subjects were asked to complete the GJT three times. The test was administered under a time constraint on the first two occasions and on the third one they completed a delayed GJT during an interview and under no time constraint. A confirmatory factor analysis (CFA) revealed two factors showing a clear distinction between those measures with time constraint (hypothesized to tap implicit knowledge) and those without any time pressure (hypothesized to tap explicit knowledge).

Following this paper, R. Ellis (2005) designed a test battery to provide independent measures of implicit and explicit knowledge which consisted of an ONT, a TGJT, an EI task, a UGJT, and an MKT. The ONT and the EI were assumed to measure implicit knowledge as participants would focus on meaning, and MKT was hypothesized to measure explicit knowledge because it requires the use of metalanguage. Regarding GJTs, the prediction was that the TGJT would tap into implicit knowledge as participants would be under time pressure, whereas the UGJT would tap into explicit knowledge because they have enough time to use their metalinguistic knowledge. Results of factor analysis showed that the ONT, TGJT, and UGJT loaded on the implicit knowledge factor, while the UGJT and MKT loaded on the explicit knowledge factor. An exploratory factor analysis (EFA) was conducted by R. Ellis, but Isemonger (2007) criticized him, based on the fact that R. Ellis tested a hypothesis based on a cognizance of the constructs and the model and should have run a CFA. R. Ellis and Loewen (2007), in response to the criticism, reanalyzed the data through CFA and verified the two-factor model from the original EFA. Then, Kachinske and Vafaee (2014) examined a one-factor model as an alternative to R. Ellis and Loewen's model, and their results demonstrated that both one-factor and two-factor models fit the data.

Therefore, the above-mentioned studies have shown that explicit and implicit knowledge are distinct constructs that can be measured with different tests and that time pressure can distinguish between them. However, some other studies evinced that time pressure does not guarantee the retrieval of implicit knowledge (DeKeyser, 2003, 2009; Gutierrez, 2013). Thus, more effort should be made to design more fine-grained measures for both knowledge types. Also, there are some other studies that validate different implicit and explicit measures, some of which, germane to the measures used in this study, will be elaborated on in the following.

Elicited imitation task

A consensus exists among L2 researchers that EI taps into linguistic knowledge. However, what resources EI exactly taps into is controversial. R. Ellis (2005), Bowles (2011), and Zhang (2015) reached the conclusion that EI loaded heavily on the factor of implicit knowledge. Erlam (2006) further indicated that the time-pressured tests highly correlated with the EI. Nonetheless, Suzuki and DeKeyser (2015) challenged the validity of EI as a measure of implicit knowledge by using EI with a Built-in WMT. They also employed an MKT as a measure of explicit knowledge. Their results showed that the performance on EI was related to MKT scores. However, their study has major limitations. First, the instructions explicitly required the participants to correct ungrammatical sentences, and this might have attracted the participants' attention to form rather than meaning. Second, they employed EI and WMT as just one task, not separately.

Grammaticality judgment test

GJTs have been extensively used to reflect L2 speakers' grammar knowledge (Qureshi, 2020). It was hypothesized that time pressure can determine the kind of knowledge GJTs tap into (Bowles, 2011; R. Ellis, 2005; Gutiérrez, 2013; Zhang, 2015). Time pressure encourages test takers to answer on the basis of their implicit knowledge, while unlimited time may let them rely on their explicit knowledge. In addition to exploring the effect of time pressure, some researchers have pointed out that task stimulus influences the constructs GJTs tap into (Gutiérrez, 2013; Loewen, 2009; Zhang, 2015). Loewen (2009) showed that there was no significant difference between the scores of the grammatical items and those of the

Vol. 19, No. 1, Spring 2022, 17-34

ungrammatical ones in UGJT. On the contrary, the ungrammatical items were significantly more difficult than the grammatical items in the TGJT. Gutiérrez (2013) also administered timed and untimed GJTs, and an MKT to 49 L2 speakers, and calculated the scores for grammatical and ungrammatical items separately. The grammatical/ungrammatical model and the timed/untimed model were compared using CFA. In the first model, both the grammatical and ungrammatical items of TGJT loaded on the construct of implicit knowledge, and the grammatical and ungrammatical items of UGJT and MKT loaded on the construct of explicit knowledge. In the second model, the grammatical items of both timed and untimed GJTs loaded on the implicit knowledge construct, and the ungrammatical items of both GJTs and MKT loaded on explicit knowledge. Results demonstrated that the first model was a better fit for the data, i.e., irrespective of time pressure, stimulus type is what distinguishes between implicit and explicit knowledge.

More recently, Vafaee et al. (2017) also examined the construct validity of GJTs by employing two new psycholinguistic measures of implicit knowledge: a self-paced reading task (SPRT) and a WMT besides previously used tasks (i.e., TGJT, UGJT, and MKT). They hypothesized that as GJTs draw participants' attention to form, manipulating time condition or stimulus type does not make them measures of implicit knowledge. Seventy-nine L2 learners participated in this study. The results confirmed their hypothesis and revealed that controlling time condition or the grammaticality of the sentences does not lead GJTs to be distinct measures of implicit and explicit knowledge.

Metalinguistic knowledge test

As metalinguistic knowledge requires deliberate attentional focus, learners will be aware when they are drawing on it. According to R. Ellis (2005), MKT is of two versions. In the first version, the problematic part of the item containing the error is underlined. The participants have to choose from among the provided options the one which best explains the error. In the other one, the participants are required to explain why the sentence is ungrammatical. The second one seems to be more valid as it makes the participants use their metalanguage without any help provided in the options. This version of MKT has been extensively used in the literature, and it has been considered to be a measure of explicit knowledge (Bowles, 2011; Gutiérrez, 2013; Kwon, 2018; Vafaee et al., 2017; Zhang, 2015).

Word monitoring task

WMT is an online measure of implicit knowledge which provides reaction time (RT) data and examines whether L2 speakers are sensitive to grammatical errors while they are reading for the purpose of comprehension. Different studies employed WMT to examine real-time sentence processing, and scrutinize the validity of the existing implicit knowledge measures (e.g., Suzuki & DeKeyser, 2015; Vafaee et al., 2017). The results of Vafaee et al. (2017) revealed that RTs to the target word in ungrammatical items were statistically greater than RTs in the grammatical items. Suzuki and DeKeyser (2017) designed a battery of tests for implicit and automatized explicit knowledge. In fact, they developed three online tasks as the tests of implicit knowledge (i.e., an SPRT, a WMT, and an eye-tracking while listening task). Then, they compared the performance of the participants in these three tasks with their performance in time-pressured tasks (i.e., timed visual/auditory GJTs and a timed fill-in-the-blanks test), used to measure automatized explicit knowledge. The results from CFA revealed that the three real-time measures tapped into the implicit knowledge factor, which was different from an automatized explicit knowledge factor as this factor was correlated with the three time-pressured tasks.

Structure Difficulty in Relation to Implicit and Explicit Knowledge

So far, the literature indicates that variation in learners' performance on different tasks can be attributed to task design features (i.e., time pressure and/or task stimulus) (Bowles, 2011; R. Ellis, 2005; Gutiérrez, 2013; Han & Ellis, 1998; Zhang, 2015). A few other studies (Bialystok, 1979; Shiu et al.,

2018) have hypothesized that variation in learners' performance may be related to the types of target structures tested. Bialystok (1979) investigated the role of structure difficulty and time constraints in learners' performance on GJTs. The GJTs were in the aural format and had two different versions based on time limit. The first version asked the learners to judge the grammaticality of the sentences within 3 seconds, while the second version allowed 15 seconds to judge the sentences. The GJTs consisted of grammatical and ungrammatical French sentences, targeting three form classes of adjectives, object pronouns, and verbs. Each form class consisted of three governing rules, classified into easy, middle, and difficult structures. It should be noted that the classification of the level of difficulty was subjective and based on the researchers' and native speakers' judgments. The results showed that the structure difficulty had a significant role in learners' performance on the 15-second GJT, whereas it did not affect learners' judgments on the 3-second GJT. In addition, learners performed better on the easy items than on the middle and difficult ones when they had more time (i.e., 15-second GJT).

A more recent study, Shiu et al. (2018) explored the role of structure difficulty as well as modality, time pressure, and task stimulus in learners' performance on GJTs. To do this, 120 EFL learners were required to judge the grammaticality of the items on timed aural, untimed aural, timed written, and untimed written GJTs. The items consisted of an equal number of grammatical and ungrammatical sentences as well as easy (i.e., past progressive) and difficult (i.e., passive voice) items. The results demonstrated that (1) modality, time pressure, and task stimulus played an important role in learners' GJT performance and (2) the difference between the performance on the easy structure and that on the difficult one was not significant although the learners performed better on the easy items. This study, however, suffers from two shortcomings that deserve attention. One important limitation of Shiu et al. (2018) is that the structure difficulty factor was limited to only two structures. Another limitation of their study is that they investigated the role of structure difficulty in GJTs without the inclusion of other measures of linguistic knowledge (e.g., WMT, EI, or MKT). This is considered a limitation as previous studies conducted CFAs in order to examine the effect of other factors such as time pressure on learners' performance. Therefore, employing several measures of linguistic knowledge and various structures is required to investigate the effect structure difficulty may have on learners' performance. To the best of our knowledge, no study has considered these two points with regard to the structure difficulty factor.

The Present Study

The present study aimed to find out whether the different tests in this study (WMT, EI, TGJT, UGJT, MKT) with different difficulty levels of structures tap into distinct constructs of knowledge or the same construct. This will let us know if item difficulty can determine the type of knowledge tapped. It further aimed to find out whether or not performance on easy/difficult items across the target structures in all the measures differs significantly. To this end, the following research questions were investigated.

1. Do WMT, EI, TGJT, UGJT, and MKT provide separate measures of implicit and explicit knowledge?
2. Do structures with different difficulty levels in WMT, EI, TGJT, UGJT, and tap into distinct constructs of knowledge?
3. Does performance on easy vs. difficult items in WMT, EI, TGJT, UGJT, MKT significantly differ?

Method

Participants

We recruited 117 learners of English as an L2 from three state universities in Tehran. Attempt was made to choose participants who were proficient enough to perform the tasks. The minimum requirement was 80 out of 100 on a C-test adapted from Ishihara, et al. (2003). Based on this criterion, 83 EFL learners were selected, out of which only 52 completed all the tests. Two more participants were eliminated from the analysis due to their low scores in answering comprehension questions of the first task. The remaining 50 participants (22 females), in the age range of 18 to 26, consisted of undergraduates (n = 19), M.A. students (n = 29), and Ph.D. students (n = 2).

Fifteen more participants who were experienced EFL teachers participated in the study to judge the target structures through a questionnaire based on R. Ellis's (2005) list of structures problematic to learners. Their judgments provided evidence to choose the target structures of the study. It should be mentioned that the tests and tasks used in this study were piloted with 9 Ph.D. TEFL candidates to ensure the quality of test items and find out the required time for GJTs.

Target Structures

As stated above, 15 experienced EFL teachers participated in the study to rate on a 1 to 5 scale (with the ends labelled very easy and very difficult) the target structures presented to them in a questionnaire including a list of structures problematic to learners (R. Ellis, 2005). According to their judgments, 8 structures (4 easy and 4 difficult structures) were chosen as the target structures and 4 structures as fillers. *Verb complement*, *regular past tense*, *comparatives*, and *possessive -s* were selected as easy structures and *unreal conditional*, *resumptive pronoun*, *transitive/intransitive verbs*, and *dative alternation* were chosen as difficult structures. Grammatical and ungrammatical sentences were developed based on these structures and used in all tasks except MKT which included only the ungrammatical ones. Sample sentences for each structure are provided below.

Sample Sentences for Easy Structures

He attempted to pass his exam, but he failed. (G) Verb Complement
*Students demanded knowing the topics of their exam. (UG) Verb Complement
When man invented the motor car, life changed for everyone. (G) Regular Past Tense
*Last year, my sisters work like professional teachers. (UG) Regular Past Tense
This old building is bigger than your house. (G) Comparatives
*It is more harder to learn Japanese than to learn English. (UG) Comparatives
His mother's face is very kind and friendly. (G) Possessive
*John lost his friend shoes yesterday in a shop. (UG) Possessive

Sample Sentences for Difficult Structures

If I were taller than you, I would play basketball. (G) Unreal Conditional
*If I had won the lottery, I would build a mansion. (UG) Unreal Conditional
The man who was standing there is my friend. (G) Resumptive Pronoun
*That's the beautiful girl whom I like her. (UG) Resumptive Pronoun
She explained the news to Mary yesterday. (G) Dative Alternation
*Rosemary reported his father the bad news. (UG) Dative Alternation
Huge problems may arise after collecting data. (G) Transitive/Intransitive
*Her English vocabulary was increased a lot last year. (UG) Transitive/Intransitive

Vol. 19, No. 1, Spring 2022, 17-34

Instruments

Five measures were used including a WMT, an EI task, two GJTs (i.e., timed and untimed), and an MKT. Moreover, to check English proficiency of the participants, a C-test was used. At the end, a background questionnaire was administered to collect demographic information about the participants.

Test of general English proficiency

A C-test including 3 passages, each with 112 to 143 words, adapted from Ishihara et al. (2003) was used. In these passages, the last half of every fifth or sixth word is omitted for the test takers to complete. There were 50 partially filled blanks. Each correct answer received two points (the maximum possible score of 100). C-test was selected because it had practical benefits for this study (10-15 minutes to administer). The reliability of C-test, using Cronbach's alpha, was .71 and scores ranged from 83 to 100 (*M* = 89.96, *SD* = .46).

Word monitoring task

The WMT was employed to measure the participants' sensitivity to grammatical errors. Following Vafaee et al. (2017) and Rezaei and Mehraein (2019), participants were shown a monitoring word in the center of the screen for 2 seconds and then a sentence appeared on the screen chunk by chunk. They were required to press a key immediately after they identified the monitoring word in the sentence. This word appeared after the related target structure (critical region) in a sentence. The first chunk of a sentence appeared on the left-hand side of the screen and after 2 seconds, the next chunk appeared automatically to the right of the preceding chunk cumulatively. Each sentence consisted of four chunks and 8-12 words. After each sentence, a true/false comprehension question was shown on the screen. In this way, participants' attention was drawn to the sentence meaning as well as the monitoring word rather than form. These comprehension questions remained on the screen for 5 seconds and participants were asked to press two fixed keys on the keyboard (left shift for false sentences and right shift for true ones). This dual-task paradigm is deemed to minimize the application of explicit knowledge. Twenty-four items composed of an equal number of grammatical, ungrammatical, easy, and difficult items were included in this task. Of these 24, 8 were fillers. Before the task, the participants were given 4 practice items to become familiar with the nature of the test. This task was delivered through DMDX (Forster & Forster, 2003). The reliability of this measure, as well as the easy and the difficult items which were calculated separately, estimated by Split-half reliability with Spearman-Brown correction were .83, .79, and .88, respectively. The following provides an example of critical region and monitoring word for *unreal conditional* structure.

Sample Sentence: *If I had more money, I <u>travel</u> <u>around</u> the whole world.
Critical Region *Monitoring Word*
Comprehension Question: I don't have enough money to travel. (True)

Oral elicited imitation ask

The EI task required the participants to (a) process an auditory stimulus sentence and (b) repeat it. This task was made up of 16 experimental sentences targeting the selected structures, half grammatical and half ungrammatical. The number of easy and difficult items was also the same. Eight fillers, 4 grammatical and 4 ungrammatical were included in this task. The sentences were aurally presented to the participants. They were told to check on an answer sheet whether they concurred with, did not concur with, or had no idea about the content of each sentence. Including the belief statements was to increase the probability of participants' focus on meaning rather than form. The statements were read and recorded by a Ph.D. candidate and an audio file was created. The participants were required to repeat the items orally. Answering belief statements and imitating the sentence had to occur within a 10-second time limit based on the results of the pilot study. This task was delivered via AIM Player and recorded using the Audacity software. The reliability of this measure and the easy and difficult items using Cronbach's alpha were .67, .69, and .71, respectively.

Timed grammaticality judgment test

This test was a computerized test including 48 sentences. Thirty-two items were presented as experimental sentences testing target structures, with an equal number of grammatical, ungrammatical, easy, and difficult structures. and 16 filler sentences were also included. This test required participants to show, in a certain period of time, if each sentence was grammatically correct or not. They were asked to utter their judgments aloud (i.e., correct/incorrect) and the experimenter would write them on an answer sheet. In this way, participants' attention was directed to form. Each sentence remained on the screen for 4.29-6.20 seconds, depending on the sentence length and the results of the pilot study, followed by a beep sound. Four practice items were also included in this task. Two lists containing the same items presented in reverse orders were created for GJTs and each participant received one of the lists. The criterion for determining the time limit for TGJT was based on the performance of a group of Ph.D. students of TEFL. Following R. Ellis (2006), mean RT for each item was calculated and then an additional 20 percent was added. This task was delivered through automated PowerPoint slides. This test's internal consistency coefficients were reported to be .70, .68, and .72 for the whole test, the easy items, and the difficult ones.

Untimed grammaticality judgment test

Untimed GJT followed a procedure like that of TGJT with the exception that there was no time constraint. This test was again computer-delivered in written modality to prevent the participants from revising their answers. Two lists of UGJT were developed which differed only in the order of items. Participants were asked to decide whether the sentences were correct or incorrect and they were reminded that there was no time constraint. The reliability estimates of the test, that of its easy items and the difficult ones were .74, .73, and .79.

Metalinguistic knowledge test

An MKT was administered to see whether participants could correct and explain the rule that the sentence was violating. Participants were presented with 16 ungrammatical sentences with two sentences for each target structure. They were said that all of the sentences were ungrammatical and there was no time limitation. This task allowed participants to access their explicit knowledge. The reliability coefficients were .81, .79, and .86 for the test, its easy items, and its difficult ones, respectively.

The features of the five measures are summarized in Table 1.

TABLE 1
Task Features of the Instruments

	WMT	EI	TGJT	UGJT	MKT
Time Pressure	Yes	Yes	Yes	No	No
Focus	Meaning	Meaning	Form	Form	Form
Modality	Written	Aural	Written	Written	Written
Data	RT	Repetition	Accuracy	Accuracy	Accuracy
Number of Items	24	24	48	48	16

It should be mentioned that the number of items in the measures was not equal. In fact, the number of items was reduced in measures hypothesized to tap into implicit knowledge owing to the nature and form

Vol. 19, No. 1, Spring 2022, 17-34

of the WMT and EI which made them more difficult tests. In these two tasks, only 16 experimental sentences and 8 fillers were included. In both GJTs, there were 32 experimental sentences and 16 fillers, and MKT consisted of 16 items. In all tasks, the ratio of grammatical, ungrammatical, easy, and difficult items was the same. However, MKT only consisted of an equal number for easy and difficult items because all items were ungrammatical.

Procedure

The five linguistic tests were administered, starting with those hypothesized to measure more implicit knowledge and progressing to tests hypothesized to measure more explicit knowledge. All the measures were administered in approximately 1.5 hours and individually. At the end of the study, the participants completed a background questionnaire. In order to reduce fatigue, there was a short break (5 min) between the first three tasks and the remaining two ones. The time and order schedule for each of the measures are shown in Table 2.

TABLE 2
Time and Order of the Instruments

Block 1		Block 2	
Task	Time	Task	Time
WMT	10 minutes	UGJT	20 minutes
EI	10 minutes	MKT	15 minutes
TGJT	10 minutes	Background questionnaire	5 minutes

Data Analysis

Before going to data analysis, we will elaborate on the scoring of the different tasks used in the study.

Word Monitoring Task: Comprehension questions and their accuracy were checked first in order to ensure that the participants paid attention to the test. The mean accuracy score was 83.63 (SD = 3.22). Two participants whose error rate was higher than 25% were eliminated from data analysis. The difference in the RT to the monitoring word between grammatical and ungrammatical sentences, and also between easy and difficult sentences provided the index for sensitivity to grammatical errors.

Elicited Imitation Task: The respondents' answers to belief statements were checked first. None of the participants had unacceptable responses. Each correctly imitated sentence was allocated a score of 1 and each incorrectly imitated one a score of 0. The maximum score for the whole task was 16 and the maximum score for grammatical, ungrammatical, easy, and difficult items was 8.

Timed and Untimed Grammaticality Judgment Test: One point was allocated to correct answers and 0 point to incorrect ones. Responses to fillers were excluded from the analyses. Thus, the maximum score for the whole test was 32. The scores were also calculated separately for grammatical and ungrammatical sentences (i.e., maximum 16), and for easy and difficult structures (i.e., maximum 16).

Metalinguistic Knowledge Test: A rubric was developed and used for scoring the responses. Partial credit was allocated to each response (i.e., one point for correct explanation, and one point for correction). Therefore, the maximum possible score was 32. In addition to one of the experimenters, one Ph.D. TEFL candidate used the rubric and rated 10 papers randomly. Their ratings were almost similar except for 3 responses. For these responses, they resolved the disagreements through discussion.

As for the analysis of the data, paired-samples t-tests were run to compare the RTs for easy/difficult and grammatical/ungrammatical items across the target structures in the WMT. Afterward, in order to examine the relationships among all the measures of the current study, Pearson-product moment correlation analyses were done before conducting CFAs. It is worth mentioning that CFA was used as the present study aimed to test the explanatory power of different existing theories and test their power through factor analysis. This is clearly a case where CFA is more appropriate than EFA.

CFAs were run using Mplus 7.11 (Muthén & Muthén, 1998-2010), testing the previously proposed models such as the two-factor implicit/explicit model reported in R. Ellis (2005) and the Grammatical/Ungrammatical model proposed in Gutiérrez (2013). Three more models were constructed and tested. The first model hypothesized that all measures tap into one construct which is linguistic knowledge. The second model hypothesized that easy items tap into automatized knowledge and difficult ones tap into explicit knowledge. In order to show evidence for the role of automatized knowledge, only GJTs were considered as appropriate measures as they have been shown to be good measures of this kind of knowledge (Suzuki & DeKeyser, 2017; Vafaee et al., 2017). The WMT was not included as it is the only psycholinguistic measure used in this study and it is recognized as a measure of implicit knowledge (Suzuki & DeKeyser, 2015; Vafaee et al., 2017). The easy items of MKT were also excluded from the model because MKT has been shown to be the purest measure of explicit knowledge and it cannot be an acceptable measure of implicit or automatized knowledge. The third hypothesized model was a three-factor model. In this model, easy and difficult items of WMT and easy items of EI were considered as measures of implicit knowledge, the easy items of both GJTs as measures of automatized knowledge, and difficult items of EI, GJTs, and MKT as measures of explicit knowledge.

Results

Descriptive Statistics

As it is evident in Table 3, regarding the easy/difficult distinction, it can be realized that the participants scored considerably higher on easy items of all tests and lower on difficult items of the tests. Additionally, the standard deviations show greater variance in the difficult items than in the easy ones.

TABLE 3
Descriptive Statistics for all the Tests Used in the Study

	Minimum	Maximum	Mean	SD	Sig (2-tailed)	*Eta-squared*
WMT	414.68	997.11	652.25	135.39		
WMT-E	344.24	1018.35	622.73	146.97	.021	.107
WMT-D	366.36	1023.93	681.77	172.59		
WMT-G	338.57	1008.56	605.25	158.14	.001	.223
WMT-U	437.53	1058.53	699.25	165.23		
EI	8	16	12.52	1.83		
EI-E	5	8	6.87	.841	.000	.487
EI-D	3	8	5.62	1.33		
EI-G	5	8	7.10	.80	.000	.737
EI-U	2	8	5.39	1.23		
TGJT	18	32	25.41	3.12		
TGJT-E	9	16	13.31	1.72	.000	.261
TGJT-D	8	16	12.16	1.98		
TGJT-G	10	16	13.85	1.45	.000	.384
TGJT-U	4	16	11.62	2.63		
UGJT	20	30	26.33	2.48		
UGJT-E	11	16	13.45	1.27	.041	.08
UGJT-D	8	15	12.91	1.76		
UGJT-G	8	16	13.29	1.73	.594	.006
UGJT-U	7	16	13.10	1.75		
MKT	18	32	28.10	3.56		
MKT-E	12	16	15.29	1.11	.000	.436
MKT-D	7	16	13.02	2.72		

Correlations

Table 4 summarizes the results of the correlation analysis. As represented in Table 4, EI correlated with TGJT, UGJT, and MKT (.28, .37, and .32, respectively) to a statistically significant degree. The TGJT correlated strongly with the UGJT (.68) and with MKT (.50). The UGJT correlated not only with the TGJT (.68) but also with the MKT (.56). MKT correlated with all measures except WMT. Finally, the WMT did not correlate with any other measures.

TABLE 4
Correlational Matrix for the Measures

	EI	TGJT	UGJT	MKT	WMT
EI		.285*	.375**	.321*	.000
TGJT			.687**	.500**	-.162
UGJT				.565**	-.107
MKT					-.107

* Correlation is significant at the .05 level (2-tailed).
** Correlation is significant at the .01 level (2-tailed).

Table 5 presents the correlation coefficients between easy and difficult structures of both GJTs and difficult items of EI and MKT. As the table shows, the correlation between all the different measures is significant excluding that of EI-D and TGJT-D (.23), EI-D and UGJT-E (.11), and MKT-D and UGJT-E (.25).

TABLE 5
Correlational Matrix for Easy and Difficult Structures in the Measures

	EI-D	TGJT-D	UGJT-D	MKT-D	TGJT-E	UGJT-E
EI-D		.234	.396**	.430**	.330*	.114
TGJT-D			.412**	.330*	.457**	.567**
UGJT-D				.449**	.616**	.363*
MKT-D					.315*	.251
TGJT-E						.384**
UGJT-E						

* Correlation is significant at the .05 level (2-tailed).
** Correlation is significant at the .01 level (2-tailed).

Research Questions 1 and 2

CFA was used to examine the adequacy of the models. The evaluation of the fit of each model was based on an inspection of the plausibility of parameter estimates and the overall fit indices provided in Mplus. In each case, attempt was made to ensure that there were no out of bound or implausible estimates. After ensuring the plausibility of individual parameter estimates, overall fit indices were examined. Mplus provides fit indices: (a) Model chi-square, (b) Root Mean Square Error of Approximation (RMSEA), (c) Comparative Fit Index (CFI), and (d) Standardized Root Mean Square Residual (SRMR).

Mplus also reports the Tucker-Lewis index (TLI) but Kline (2015) warns against the use of this index as it is highly correlated with CFI. The model chi-square tests the "exact-fit" hypothesis for the model. Since almost all models are expected to diverge from the data to some extent, the chi-square usually indicates poor fit in almost all cases.

It is usually recommended (e.g., Brown, 2006) that RMSEA values be less than .06 to indicate good fit. Usually, a 90% interval is also reported for the RMSEA. The upper bound of this interval should not be higher than .08. Moreover, the CFI must be higher than .90 and the SRMR should not exceed .08. The CFA analyses in Mplus revealed that out of the initial five hypothesized models, three models did not converge to an acceptable solution. The problem was resolved through increasing the number of iterations. The remaining two models showed acceptable fit.

The Journal of Asia TEFL
Vol. 19, No. 1, Spring 2022, 17-34

Table 6 shows the fit indices for the two models with adequate fit. Moreover, an inspection of the individual parameter estimates revealed that all of them were logical and there were no out of bound estimates. Hence, it may be concluded the one-factor model and the easy/difficult model adequately fit the data.

TABLE 6
Fit Indices for the Initial CFA Models

	Chi-square	RMSEA		CFI	SRMR
		Index	90 % interval		
One-factor model	1.070	.00	.00-.00	1.00	.023
Easy/Difficult model	6.229	.00	.00-.162	1.00	.043

The path diagrams along with the related parameter estimates for the two models are displayed in Figures 1 and 2. The evaluation of the fit of each model was based on an inspection of the plausibility of parameter estimates and the overall fit indices. All estimates were plausible. In the one-factor model, all tests in this study loaded on one factor. This factor can be referred to as linguistic knowledge. It should be mentioned that although the factor loading for WMT was .14, still it was significant.

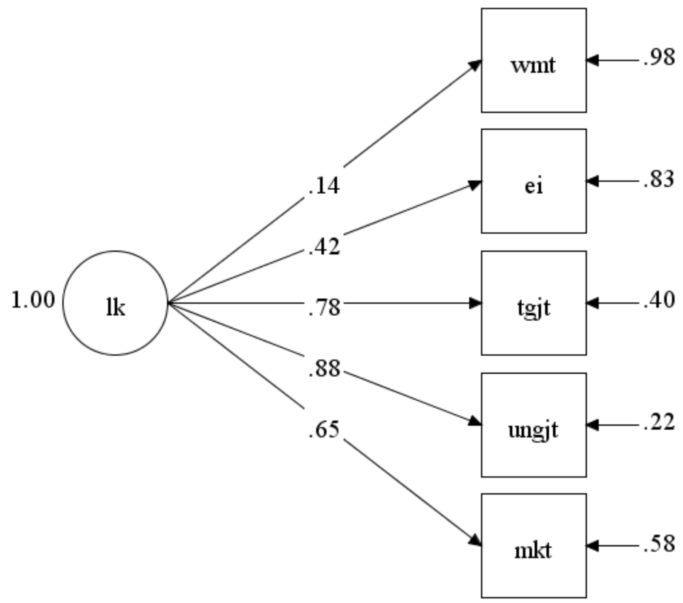

Figure 1. The one-factor model.
Note. lk = linguistic knowledge; wmt = word monitoring task; ei = elicited imitation; tgjt = timed grammaticality judgment test; ungjt = untimed grammaticality judgment test; mkt = metalinguistic knowledge test.

In the second model which was the easy/difficult model, structures with different difficulty levels tapped into distinct constructs of knowledge. Specifically, the easy items of GJT, irrespective of time condition, tapped into one construct and the difficult ones into another construct. The former was named automatized knowledge; it could not be named implicit knowledge because UGJT easy items, including grammatical and ungrammatical ones, loaded on this construct and evidence in the literature suggests that ungrammatical items in a UGJT cannot be a measure of implicit knowledge. The latter construct was named explicit knowledge because MKT, known as the pure measure of explicit knowledge, loaded on it (Figure 2).

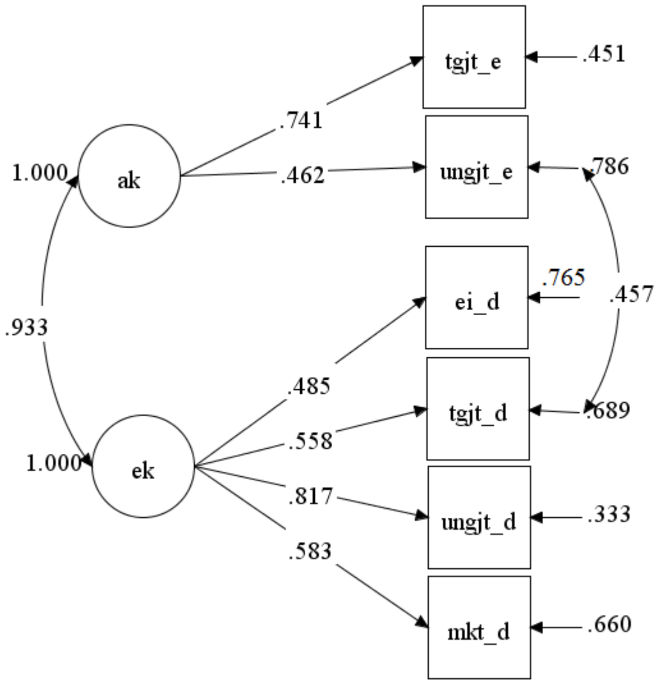

Figure 2. The easy/difficult model.

Note. ak = automatized knowledge; ek = explicit knowledge; tgjt_e = timed grammaticality judgment test, easy items; ungjt_e = untimed grammaticality judgment test, easy items; ei_d = elicited imitation, difficult items; tgjt_d = timed grammaticality judgment test, difficult items; ungjt_d = untimed grammaticality judgment test, difficult items; mkt_d = metalinguistic knowledge test, difficult items.

As can be seen in Figure 2, only the easy items of GJTs were included in the model as measures of automatized knowledge. This is due to the fact that WMT, as the only pure measure of implicit knowledge, did not correlate with any other measure.

As shown in Figure 2, there is a significant correlation between the difficult items of TGJT and the easy items of UGJT (.457). A plausible interpretation of this could be that firstly the nature of these two tests is similar in that both require the participants to judge the grammaticality of the sentences. Secondly, the participants completed the TGJT, which imposed a time constraint, followed by UGJT. Thus, judging difficult items under time pressure might have helped the learners to judge the easy items of UGJT more quickly, although there was not any time constraint.

Apart from these two fitted models, three rival models were also tested, two based on the findings of previous studies, and one model developed in the present study. The first rival model was based on the difference between timed and untimed tests. This model did not converge to an acceptable solution. The second rival model, following Gutiérrez (2013), was based on the distinction between grammatical and ungrammatical sections. This model did not converge to an acceptable fit either. An inspection of the data and software warnings revealed that there are serious problems in the specification of the models which could not be resolved with minor model revisions. Another rival model including implicit, automatized, and explicit knowledge was hypothesized in the study as an alternative model. In this three-factor model, WMT and easy items of EI were considered as measures of implicit knowledge, the easy items of both GJTs as measures of automatized knowledge, and difficult items of EI, GJTs, and MKT as measures of explicit knowledge. However, this model did not show an acceptable fit for the data of this study.

Vol. 19, No. 1, Spring 2022, 17-34

Research Question 3

Paired-samples t-tests were conducted to compare the performance on easy/difficult items across the target structures in all the measures. As Table 4 indicates, significant differences exist between the performance on the easy items and that on the difficult ones in all the measures. Four other paired-samples t-tests were conducted to compare learners' performance on grammatical/ungrammatical items in all the measures. This was done in order to check the role of task stimulus in learners' performance in the current study as this factor was shown to play an effective role (e.g., Gutiérrez, 2013; Vafaee et al., 2017). The results showed significant effects of difficulty and grammaticality in all the tasks except UGJT (see Table 4). As shown in Table 4, the squared statistic in all cases indicated a large effect size excluding UGJT where easy vs. difficult items had only a moderate effect and grammatical vs. ungrammatical items had a small effect. Altogether, these findings indicate that both task stimulus and structure difficulty play pivotal roles in learners' performance.

Discussion

There is no consensus regarding tests that can measure learners' implicit and explicit knowledge. Previous studies revealed that time condition or stimulus type are factors which convert different tests into measures of different types of knowledge (Bowles, 2011; R. Ellis, 2005; Gutiérrez, 2013; Han & Ellis 1998; Zhang, 2015). Yet, the role of structure difficulty in relation to these two knowledge types has received less attention (Bialystok, 1979; Shiu et al., 2018). Furthermore, few empirical studies have provided evidence for the difference between implicit and automatized knowledge (Suzuki & DeKeyser, 2015, 2017; Vafaee et al., 2017). However, indicating this difference and examining the role of structure difficulty by comparing different measures of knowledge in CFA models is absent in previous studies. The current study tried to (1) find out whether different tests in this study tap into distinct constructs of knowledge or the same construct and whether structures with different levels of difficulty tap into different constructs of knowledge, and (2) see whether there are differences between easy/difficult structures in all the measures.

Concerning the first goal, alternative CFAs were carried out to examine whether the scores of the participants could fit a model. Two of these models were based on the findings of previous studies, i.e., timed/untimed and grammatical/ungrammatical. These two did not converge to an acceptable solution. Regarding the three models hypothesized in this study, results showed that all tasks loaded on one factor. This indicates that the participants used their linguistic knowledge to perform the tasks. The second CFA demonstrated that the easy items loaded on one factor and the difficult ones on another factor. The first construct including easy items of timed and untimed GJTs was named automatized knowledge but not implicit knowledge as it has been consistently shown that UGJT cannot be a measure of implicit knowledge. In fact, accessing explicit knowledge, even in a fast condition, requires awareness (i.e., automatized knowledge), which is different from implicit knowledge. The second construct including the difficult items of EI, TGJT, UGJT, and MKT is explicit knowledge as MKT has been recognized as a pure measure of explicit knowledge. This finding suggests that structure difficulty is also a factor in discriminating automatized and explicit knowledge. However, a high correlation between the two hypothesized factors (.93) was found in the fitted model; therefore, it is hard to claim that the two factors are totally distinct. Rather, the two share the same nature, as neither has become implicit. The only difference between them is that one has become automatized. Another model including three factors of implicit, automatized, and explicit knowledge was hypothesized in the study as a rival model. In this three-factor model, WMT and easy items of EI were considered as measures of implicit knowledge, the easy items of both GJTs as measures of automatized knowledge, and difficult items of EI, GJTs, and MKT as measures of explicit knowledge. However, this model did not show an acceptable fit for the data of this study. Generally, in any model at least two measures per factor are needed: "Two might be fine, three is better, four is best, and anything more is gravy" (Kline, 2015, p. 195); the present study included one measure of implicit knowledge (i.e., WMT), two automatized knowledge measures (i.e., the easy items of TGJT and UGJT), and four explicit knowledge measures (i.e., the difficult items of EI, TGJT, UGJT, and MKT).

Regarding the second goal of this study which was about the performance differences between easy/difficult structures, the results of t-tests on different types of items in all the tasks showed that there were significant differences between learners' performance on easy and difficult items. This can be due to the fact that the difficulty level of structures affected the participants' sensitivity to grammatical errors in psycholinguistic tasks such as WMT and/or the accuracy of their scores in other tasks. Furthermore, the participants' performance on the grammatical items and ungrammatical items differs significantly in all the tasks except UGJT. Therefore, these results show that both task stimulus and structure difficulty play a key role in the participants' online sensitivity to grammatical errors.

Conclusion

The body of literature has considered the role of time condition and task stimulus, but little research has empirically investigated the role of structure difficulty. Moreover, little attention has been devoted to the difference between implicit knowledge and automatized knowledge. The present study addressed these gaps, confirmed the claim regarding the role of structure difficulty, and provided evidence for the existence of automatized knowledge. The findings have further enriched our understanding of learners' online sensitivity to grammatical errors in a psycholinguistic measure such as WMT and/or their ability to accurately respond to test items. In fact, the results of this study indicated that there are significant differences between easy/difficult and grammatical/ungrammatical sentences in (psycho)linguistic tasks. Therefore, the results showed that structure difficulty significantly affects the L2 learners' performance. The present findings demonstrated that performance on GJTs can reliably reflect the use of automatized knowledge and suggested that even under time pressure, L2 learners can access automatized knowledge.

This study has provided evidence for the role of structure difficulty in identifying different knowledge types. Obviously, future research is needed regarding the effects of structure difficulty on the learners' performance on different tests. Moreover, this study highlights the need for examining the scores on easy/difficult and grammatical/ungrammatical items in other psycholinguistic tests such as SPRT. Moreover, another variable, WMC, may have influenced the participants' scores. In fact, learners who have high WMC may perform well on the WMT and EI.

The Authors

Sepideh Mehraein is currently a Ph.D. candidate in Teaching English as a Foreign Language at the University of Tehran, Iran. She has been teaching English at different institutes and universities. Her main research interests include second language acquisition, psycholinguistics, and teacher education. She has presented papers and workshops in national conferences. Her publications include papers in *Iranian Journal of Applied Linguistics* and *Cypriot Journal of Educational Sciences.*

Department of English Language and Literature
Faculty of Foreign Languages and Literatures
University of Tehran
Karegar-e-Shomali st. adjacent to the Faculty of Physical Education, Tehran, Iran.
Email: sepideh.mehraein@ut.ac.ir

Hamideh Marefat is a Professor in Applied Linguistics at the University of Tehran, Iran. Her research focuses mainly on second language acquisition, psycholinguistics, and processing sentences in a second language. She has been teaching Linguistics, Second Language Acquisition, and Psycholinguistics to graduate students. Her publications include papers in *Computer-Assisted Language Learning, European Journal of Developmental Psychology, Journal of Cognitive Science, Iranian Journal of Applied Linguistics, Journal of Child Language, Journal of Information Technology Education: Research,* and *Studia Linguistica.*

Department of English Language and Literature
Faculty of Foreign Languages and Literatures
University of Tehran
Karegar-e-Shomali st. adjacent to the Faculty of Physical Education, Tehran, Iran.
Email: marefat@ut.ac.ir

Hossein Karami is an Assistant Professor of Applied Linguistics/TESOL at the English Department of the University of Tehran, Iran. His areas of interest include validity and fairness, especially in the context of language testing. His research has been published in various international scholarly journals including *Language Testing, International Journal of Bilingual Education and Bilingualism, Educational Research and Evaluation, RELC Journal, Psychological Test and Assessment Modeling, TESOL Journal, Asia-Pacific Education Review,* and *International Journal of Language Studies.*

Department of English Language and Literature
Faculty of Foreign Languages and Literatures
University of Tehran
Karegar-e-Shomali st. adjacent to the Faculty of Physical Education, Tehran, Iran.
Email: hkarami@ut.ac.ir

References

Anderson, J. R. (2005). *Cognitive psychology and its implications* (6th ed.). Worth Publishers.

Barrot, J. S. (2020). Revisiting the interface positions in second language acquisition: Towards a continuum-interface model. *The Journal of Asia TEFL*, *17*(2), 616-625. https://dx.doi.org/10.18823/asiatefl.2020.17.2.19.616

Bialystok, E. (1979). Explicit and implicit judgements of L2 grammaticality. *Language Learning*, *29*(1), 81-103. https://doi.org/10.1111/j.1467-1770.1979.tb01053.x

Bowles, M. A. (2011). Measuring implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition*, *33*(2), 247-271. https://doi.org/10.1017/S0272263110000756

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. Guilford Press.

DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & H. M. Long (Eds.), *The handbook of second language acquisition* (pp. 312-348). Blackwell Publishers.

DeKeyser, R. M. (2009). Cognitive-psychological processes in second language learning. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching,* (pp. 119-138). Wiley-Blackwell.

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*(2), 305-352. https://doi.org/10.1017/S027226310505014X

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. *Studies in Second Language Acquisition, 27*(2), 141-172. https://doi.org/10.1017/S0272263105050096

Ellis, R. (2006). Modelling learning difficulty and second language proficiency: The differential contributions of implicit and explicit knowledge. *Applied Linguistics*, *27*(3), 431-463.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford University Press.

Ellis, R. (2009). Measuring implicit and explicit knowledge of a second language. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching,* (pp. 31-64). Multilingual Matters.

Ellis, R., & Loewen, S. (2007). Confirming the operational definitions of explicit and implicit knowledge in Ellis (2005): Responding to Isemonger. *Studies in Second Language Acquisition, 29*(1), 119-126. https://doi.org/10.1017/S0272263107070052

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, *27*(3), 464-491. https://doi.org/10.1093/applin/aml001

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 116-124. https://doi.org/10.3758/BF03195503

Gutiérrez, X. (2013). The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. *Studies in Second Language Acquisition, 35*(3), 423-449. https://doi.org/10.1017/S0272263113000041

Han, Y., & Ellis, R. (1998). Implicit knowledge, explicit knowledge and general language proficiency. *Language Teaching Research, 2*(1), 1-23. https://doi.org/10.1177/136216889800200102

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition*, *27*(2), 129-140. https://doi.org/10.1017/S0272263105050084

Isemonger, I. M. (2007). Operational definitions of explicit and implicit knowledge: Response to R. Ellis (2005) and some recommendations for future research in this area. *Studies in Second Language Acquisition*, *29*(1), 101-118. https://doi.org/10.1017/S0272263107070040

Ishihara, K., Hiser, E., & Okada, T. (2003). Modifying C-test for practical purposes. *Doshisha Studies in Language and Culture*, *5*(4), 539-568.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513-541. https://doi.org/10.1016/0749-596X(91)90025-F

Kachinske, I., & Vafaee, P. (2014). *Reexamining the validity of GJTs as measures of implicit knowledge: Reanalysis of Ellis & Loewen (2007), Bowles (2011), and Gutierrez (2013)* [Conference presentation]. The Second Language Research Forum (SLRF), University of South Carolina, Columbia, South Carolina.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Kwon, S.-H. (2018). A study on NNS teachers' language awareness. *The Journal of Asia TEFL*, *15*(1), 51-65. https://doi.org/10.18823/asiatefl.2018.15.1.4.51

Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching,* (pp. 94-112). Multilingual Matters.

Muthén, L., & Muthén, B. O. (1998-2010). *Mplus users guide* (6th ed.). Muthén & Muthén.

Qureshi, M. A. (2020). Grammaticality judgment task: Reliability and scope. *The Journal of Asia TEFL, 17*(2), 349-362. https://dx.doi.org/10.18823/asiatefl.2020.17.2.3.349

Rezaei, A., & Mehraein, S. (2019). Implicit and explicit instruction and EFL learners' implicit knowledge development: Evidence from word monitoring task. *Iranian Journal of Applied Linguistics, 22*(1), 116-153.

Shiu, L. J., Yalçın, Ş., & Spada, N. (2018). Exploring second language learners' grammaticality judgment performance in relation to task design features. *System*, *72*, 215-225. https://doi.org/10.1016/j.system.2017.12.004

Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, *65*(4), 860-895. https://doi.org/10.1111/lang.12138

Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning*, *67*(4), 747-790. https://doi.org/10.1111/lang.12241

Vafaee, P., Suzuki, Y., & Kachisnke, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition*, *39*(1), 59-95. https://doi.org/10.1017/S0272263115000455

Zhang, R. (2015). Measuring university-level L2 learners' implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition*, *37*(3), 457-486. https://doi.org/10.1017/S0272263114000370

(Received March 30, 2021; Revised February 20, 2022; Accepted March 18, 2022)